Comparison of COVID19 policies using a SIR-model

Linus Antell

School of Science

Bachelor's thesis
Espoo 17.8.2021

Supervisor and advisor

Prof. Kai Virtanen

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the bachelor's thesis

Author Linus Antell

Title Comparison of COVID19 policies using a SIR-model

Degree programme Engineering Physics and Mathematics

Major Mathematics and Systems Sciences **Code of major** SCI3029

Teacher in charge Prof. Kai Virtanen

Advisor Prof. Kai Virtanen

Date 17.8.2021 **Number of pages** 27 **Language** English

Abstract

In this thesis, a SIR-model (susceptible-infected-recovered model) is used to compare different non-pharmaceutical intervention policies to deal with the COVID19 epidemic. In the SIR-model, the whole population is compartmentalized into susceptible, infected and recovered populations. Susceptible individuals start as healthy individuals but are at risk of becoming infected. Infected individuals can infect susceptible individuals during the infectious period. The new infections occur at a rate referred as the rate of transmission. Infected individuals become recovered individuals after the infectious period. The recoveries occur at a rate referred as the rate of recovery. The rate of transmission and recovery and the intial state of the population are the only parameters that affect the outcome of the SIR-model. This simplicity makes for pleasant fitting of the SIR-model curves to data sets. Therefore, the effects of vaccination are specifically excluded to keep the model simple. The SIR-model is fitted to data sets from four european countries using sum of least squares estimation. In addition, using a complementary data set containing a list of governerment policies of these countries a general understanding of what policies are the most advantageous is obtained.

The results of this thesis imply that the policies analyzed do reduce the rate of new infections. In particular policies that reduce physical contacts seem to have a larger impact. It is also concluded that policies should be sustained for a long time to make a macroscopic difference in total case count, as the amount of new infections seems to accelerate whenever the policies are removed.

Keywords SIR-model, COVID19, non-pharmaceutical intervention, sum of least squares estimation

Aalto University
School of Science

Aalto-universitetet, PB 11000, 00076 AALTO
www.aalto.fi
Sammandrag av kandidatarbetet

Författare Linus Antell

Titel Jämförelse av covid-19 åtgärder med hjälp en SIR-modell

Utbildningsprogram Teknisk Fysik och Matematik

Huvudämne Matematik och Systemanalys **Huvudämnets kod** SCI3029

Ansvarslärare Prof. Kai Virtanen

Handledare Prof. Kai Virtanen

Datum 17.8.2021 **Sidantal** 27 **Språk** Engelska

Sammandrag

I denna avhandling använder vi SIR-modellen (SIR från engelska 'susceptible-infected-recovered') till att jämföra olika icke-farmaceutiska åtgärder för att minska på spridningen av covid-19 som en strategi för att hindra epidemin. SIR-modellen delar hela befolkningen in i mottagliga, infekterade och återhämtade grupper. Mottagliga individer är friska i början men riskerar att bli smittade. Infekterade individer kan infektera mottagliga individer under infektionsperioden. Hur stor del av de mottagliga individerna som blir infekterade beror på den infekterade gruppens storlek samt smittotakten. Infekterade individer återhämtar sig efter infektionsperioden och blir således återhämtade individer. Enligt SIR-modellen kan dessa individer sedan inte bli infekterade flera gånger. Hur snabbt återhämtningen sker beror på återhämtningstakten. Dessa smitto- och återhämtningstakter samt befolkningens utgångstillstånd i början är de enda parametrarna som påverkar det slutgiltiga utfallet av SIR-modellen. Den låga mängden parametrar gör det enkelt att göra en anpassning av SIR-modellen till covid-19 data. Man kan dessutom vidare förenkla anpassningen genom att enbart anpassa smittotakten. Detta kan man göra med antagelsen att infektionsperioden är konstant samt om informationen i datan används som utgångstillstånd. Således utesluter vi också vaccineringens inverkan för att behålla modellens enkelhet, samt för den orsaken att vaccinering är en farmaceutisk motåtgärd.

I denna avhandling anpassar vi SIR-modellen på covid-19 data från fyra europeiska länder genom minstakvadratsmetoden. Dessutom använder vi skild data som innehåller listor över genomförda covid-19 åtgärder i dessa länder. Detta gör vi för att få en form av allmän förståelse för vilka sorters åtgärder som är mest effektiva för att hindra covid-19-epidemin. Våra resultat tyder på att dessa icke-farmaceutiska åtgärder minskar på tillväxten av nya infektioner. Särskilt åtgärder som minskar på fysiska kontakter i befolkningen tycks ha stor inverkan. Vi drar också den slutsatsen att åtgärderna bör hållas i kraft under en lång tid för att göra en skillnad i mängden totala infektioner på lång sikt. Detta eftersom tillväxten av nya infektioner verkar öka på nytt när åtgärderna tas bort.

Nyckelord SIR-modell, covid-19, icke-farmaceutiska åtgärder, minsta kvadratmetoden

Contents

1 Introduction

The COVID19 pandemic has been an inconvenience to society for over two years and is still ongoing. Many strategies have been employed to contain the spread of this disease. Governments have used NPI (non-pharmaceutical intervention) policies to stop the spread of this virus, such as lockdowns. However, these policies come with an economic trade-off (Ferguson et al., 2020), (Eichenbaum et al., 2020) and can be annoying to live with for a population. In this thesis, we assess some of the impact these policies have on the pandemic. This assessement is done by comparing the situation of the pandemic in four different countries, more specifically the progression of the case numbers is of interest.

The comparison is conducted with the help of the SIR-model (susceptible-infected-recovered) (Kermack and McKendrick, 1927), which we use for its simplicity and previous use in mathematical epidemiology (Brauer, 2017), including recent applications to COVID19 (Cooper et al., 2020). In the SIR-model, the whole population is compartmentalized into susceptible, infected and recovered populations. Infected individuals infect susceptible individuals and remain infectious during a time called the infectious period. After this period infected individuals become and remain recovered individuals. The outcome of the SIR-model is completely determined by the initial state of the population and two parameters that are the rates of recovery and transmission. In this thesis, the SIR-model is fitted using sum of least squares estimation with data from the four countries, which gives us insight on changes in the rate of transmission over time. NPI policies have been known to reduce the rate of transmission (Ferguson et al., 2020). Therefore inspecting changes to the rate of transmission allow us to compare the impact of such policies. For clarity, by NPI policy we mean any prevention strategy for COVID19 that does not involve the use pharmaceutical substances. For example, lockdowns or mask mandates are included while the effects of antiviral drugs are excluded from the analysis. In addition, the relation between the rate of transmission and the policies are compared using a government response data set from the ECDC. Though, there are some limits with such an approach because of the general structure of the response data set. Important policies may be excluded from our analysis for two main reasons. The sources used in the data set are based of information the governments have made publicly available on the internet. It is therefore possible that not every policy is listed in the data set. Furthermore, the policies in the data sets are formatted such that similar policies are listed as the same, but these policies are not always identical.

The main objective of this thesis, is to have a look on how the COVID19 situation in the chosen countries relates to the rate of transmission. We also wish to gain insight into the characteristics of policies that universally reduce the rate of transmission. The focus is NPI policy so the impact of vaccination is not considered in this thesis, as it is a pharmaceutical intervention and this would also further complicate the SIR-model.

The thesis is structured as follows: Section 2 is a brief introductory discussion of a few topics relating to mathematical epidemiology. Section 3 is a detailed explanation of the methodology used in this thesis. Section 4 presents the results using the methods from the previous section. In Section 5, the results are discussed. Section 6 presents a conclusion.

2 Background

Mathematical modeling in epidemiology can be utilized as a powerful tool that provides decision makers a way to study disease in a safe environment. Thereby granting insight into the core dynamics of infectious disease spreading in populations. An important class of epidemiological models are compartmental models, where the disease is studied in a population subsetted into compartments. A common compartmental model in mathematical epidemiology is a SIR-model (Kermack and McKendrick, 1927). Where the population is compartmentalized such that it consists of susceptible, infected and recovered individuals. The SIR-model and even more compartmental models, such as SIRD (SIR-deceased) models, have been used to accurately model previous pandemics such as the 1918 influenza epidemic, HIV or Ebola (Brauer, 2017).

Many characteristics of infectious diseases are revealed through mathematical modeling. Some of these characteristics relate to simple epidemiological metrics. Since many models can be applied to one disease, it is not too surprising that the models also relate to each other via epidemiolocial metrics. An important metric of infectious diseases that characterizes the behaviour of many models, is the basic reproduction number commonly denoted by R_0. The basic reproduction number is the expected number of new infections resulting from one infected individual and may have spatiotemporal dependance. For example, using branching processes it can be found that diseases with a R_0 value greater than one are likely to persist for a long time and may even become endemic, while diseases with R_0 less than one are likely to go extinct. The basic reproduction number relates to the SIR-model as the ratio between the rates of transmission and recovery of the SIR-model (Brauer, 2017).

Mathematical epidemiology has recently been applied to COVID19 research. In Bertozzi et al. (2020), a general understanding of the challenges of COVID19 modeling are obtained using three different models. An exponential growth model, a branching process model and a SIR-model were fit to COVID19 data. They suggest that the different models could be combined for forecasting together. The models would be combined such that one transitions from one model to another during different stages of the epidemic. A key problem in COVID19 modeling is the importance of testing rates, which are tests performed per population. Lower testing rates may increase the amount of outliers in the data, which causes volatility in parameter estimation such that the models become unreliable. A SIR-model approach, similar to what we analyze in this thesis, was used in Cooper et al. (2020) to gain insight into early dynamics of the pandemic. Another example of SIR-model application to COVID19

introduced by You et al. (2020), where the SIR-model was used in conjunciton with an exponential growth model. In You et al. (2020), the basic reproduction number in different regions of China was estimated along with other epidemiological metrics. For example the infectious period, which is the time period one remains infected, was estimated to be two weeks long. The basic reproduction number of COVID19 is highly variable and has been commonly found to be reduced by NPI policies, such as lockdowns (Ferguson et al., 2020), (Bertozzi et al., 2020), (You et al., 2020), (Cooper et al., 2020). Moreover, such policies need be sustained over a longer period of time, as epidemics return to growth phases whenever the policies are removed (Bertozzi et al., 2020).

In this thesis, a SIR model will be used to analyze the COVID19 situation in four european countries. For each of the countries we estimate the rate of transmission. Which relates to how easily new infections are transmitted. The estimated rates of transmission are then used to compare the COVID19 response in the countries. In addition, using a list of implemented policies in the countries should provide insight into general characteristics of efficient NPI policies.

3 Methods

In this thesis we fit the SIR-model to the COVID19 (case)data of Finland, France, Italy and Sweden, using done by sum of least squares estimation. In this section, the methods are discussed in detail.

3.1 SIR-model

In the SIR-model (Kermack and McKendrick, 1927), the total population is compartmentalized into three compartments, the susceptible-, infected- and recovered population. We assume that the total population remains fixed. The susceptible population represent the currently healthy, but able to be infected individuals. The infected population represents the infected individuals. Some proportion of susceptible individuals in contact with infected individuals will become infected. The recovered population consists of individuals who were previously infected but then recovered (or deceased). At each consecutive point of time, a proportion of the infected population recovers from the infection. It is assumed that an individual cannot become infected twice. It is also assumed that the population is so large that the contact mixing is homogeneous. That is, there are no specific intricate rules for who can contact whom and an infected individual is guaranteed to infect that certain fraction of susceptible individuals. Here, we include deceased individuals in the recovered population for simplicity. Lastly, since the disease is airborne we assume a density dependence, where a larger amount of possible contacts in a population increase risk of transmission. The SIR-model is thus given by the following system

of differential equations:

$$\frac{dS}{dt}(t) = -\beta \frac{S(t)I(t)}{N}, \tag{1}$$

$$\frac{dI}{dt}(t) = \beta \frac{S(t)I(t)}{N} - \gamma I(t), \tag{2}$$

$$\frac{dR}{dt}(t) = \gamma I(t), \tag{3}$$

where at any given time t, N is the total population, $S(t)$ is the susceptible population, $I(t)$ is the infected population and $R(t)$ is the recovered population. The rate of transmission β is the rate of new infections per infected in contact with a susceptible individual. The recovery rate γ is the rate at which infected individuals become recovered. Here we assume constant rate of recovery. If D is the total time during which an infected individual is able to infect susceptible individuals, the rate of recovery is given by $\gamma = \frac{1}{D}$ and by the results of You et al. (2020) D is set to two weeks (14 days).

We assume that NPI policies affect the rate of contacts an infected individual has as well as the risk of an infection per contact. Therefore, β should vary with time especially when policies change. For simplicity, we assume that β remains constant during consecutive time intervals (with varying lengths), depending on what policies are in effect during those time intervals.

As we compare populations of varying size it is reasonable to scale the total population to unity. Normalized variables are obtained by dividing each variable by the total population N, and they are denoted by lower case letters s, i, r for S, I, R, respectively. For easier computing, the differential equations are transformed into difference equations, such that each differentiation with respect to time is swapped with consequtive differences $\Delta x(t) = x(t) - x(t-1)$. Using a difference equation in lieu of a differential equation may lead to somewhat different results. This should be macroscopically insignificant, as every ordinary differential equation has a corresponding difference equation that behaves as a good enough approximation by a theorem by Mickens (1988). By macroscopic insignificance mean the general shape of the variable curves obtained using the equations as well as the size scales of the variables. For example, while similarly sized, graphically the solutions of differential equations are smoother than the solutions of the corresponding difference equations. A system of difference equations simplifies calculation, as each variable can now be calculated recursively. The SIR-model that is used here is then given by the following system of equations:

$$\Delta s(t+1) = -\beta s(t) i(t), \tag{4}$$

$$\Delta i(t+1) = \beta s(t) i(t) - \gamma i(t), \tag{5}$$

$$\Delta r(t+1) = \gamma i(t), \tag{6}$$

where s, i, r are the susceptible-, infected- and recovered proportions of the population. It should be noted that t is in weeks. This is important when comparing results from other papers, which can be done using the relation between the SIR-model

parameters and the basic reproduction number: $R_0 = \frac{\beta}{\gamma}$. For example, if $\beta = \frac{7}{2}$ when time is in weeks then $R_0 = \frac{\frac{7}{2}}{\frac{1}{2}} = 7$, as γ is constant $\gamma = \frac{1}{2}$. To convert β when time is in days the relation $\beta = R_0 \cdot \gamma$ is used: $\beta = 7 \cdot \frac{1}{14} = \frac{1}{2}$ (since $\gamma = \frac{1}{14}$ when t is in days).

3.2 Data and estimation

The two data sets that are used, are obtained from the ECDC website (https://www.ecdc.europa.eu/en). One data set contains weekly case numbers and other COVID19 related information ECDC (2021*a*). The other data set is a government response data set, which contains lists of policies enacted in countries in the EU ECDC (2021*b*). We choose to use the data of Finland, France, Italy and Sweden. The rate of transmission for each respective country is then estimated.
For estimation, the infected proportion of the population $i(t)$ needs to be determined at each time point, as it is not explicitly included in the data set. Since we assume that an infected individuals recovers after the time D, the recovered proportion of the population is the cumulative cases after lag D. Therefore, $r(t)$ is given by the following equation:

$$r(t) = \sum_{j=0}^{t-D} \Delta i(j) = c(t - D), \tag{7}$$

$$r(t) = 0, \text{ when } t < D, \tag{8}$$

where $c(t)$ is the cumulative infected proportion of the population. Where the aforementioned proportion are the cumulative cases per total population $c(t) = \frac{\sum I(t)}{N}$. The infected proportion of the population is then simply the proportion of cumulative cases substracted by the recovered proportion of the population, which is obtained from the following equation:

$$i(t) = c(t) - r(t). \tag{9}$$

We divide the total time period into smaller somewhat arbitrary consecutive time intervals. The time intervals are chosen such that the model fit becomes somewhat realistic. This is determined mostly visually based on the shape of the curves, but also such that policy changes are accounted for. In other words, whenever a lot of policies changes occur, a new time interval begins. This is also somewhat arbitrary as sometimes policies come into effect at different weeks and then they overlap. The time intervals are chosen from the time before vaccinations start, which is 2020 week 52 for the EU, as the effects of vaccination is excluded from our analysis. This leaves us with a total of 52 data points of weekly cases per country in consideration.
For the sum of least squares estimation a loss function l is used defined by the following equation:

$$l(\hat{\beta}) = \sum_t (i(t) - \hat{i}(t))^2. \tag{10}$$

The estimated infected proportion of the population is fitted using sum of least squares approximation. The objective is to minimize the loss function l (10) where

the sum is over all of the 52 entries, $i(t)$ is the known infected proportion of the population (obtained from the data) and $\hat{i}(t)$ is the fitted infectded proportion of the population. The β estimate is obtained starting with an initial guess for $\hat{\beta}$, a step is taken either in a positive or negative direction. The chosen step is determined by comparing whichever of the steps has the lowest sum of squares. If the step taken results in a larger sum of squares, the step is divided by 10. The last step is repeated many times until a step that lowers the sum of squares is obtained. This is iterated until a small enough sum of square is obtained, or if no further progress is made for many iterations. For our purposes a small enough sum of squares is 10^{-10} and 300 iterations is the highest amount of iterations. These numbers are arbitrarily chosen. Here we assume that the human eye shouldn't be able to spot the difference between any two results produced using such numbers with the above methods.
The same sum of least squares estimation procedure that was described above is given using shorter symbolic notation by the following equations:

$$\hat{\beta}^{+} = \hat{\beta}_n + d, \tag{11}$$

$$\hat{\beta}^{-} = \hat{\beta}_n - d, \tag{12}$$

$$\hat{\beta}_{n+1} = argmin_{\hat{\beta}}(\{l(\hat{\beta}^{+}), l(\hat{\beta}^{-})\}), \tag{13}$$

where d is the stepsize, $\hat{\beta}^{+}$ and $\hat{\beta}^{-}$ are the estimates used to compare the sum of squares and β_n is the n:th iteration. For each new time interval, the initial value of $\hat{i}(t_0)$ is set to the value of $i(t_0)$ at the beginning of that interval (at $t = t_0$). Estimation is then done for each interval. The estimation procedure is implemented using the R-language (code given in Appendix A).

4 Results

The countries chosen for SIR-model fitting were Finland, France, Italy and Sweden. For of each country the rate of transmission was estimated during different consecutive time intervals. The smallest time intervals were 3 weeks long while the largest were up to 12 weeks long. The intervals were chosen based of two criteria. The time intervals were chosen such that they match visual changes in the growth rate of the curves. In addition, the time intervals were also chosen such that as little policy change occurs within the time intervals. The earliest time points start when the infected population is larger than one. In other words, the period before the first recorded infections is skipped. In Table 1 the estimated values of the rate of transmission are presented. There should be caution when comparing $\hat{\beta}$ between the countries as the intervals vary by country.

Table 1: Estimated rate of transmission $\hat{\beta}$ of different countries at varying time intervals

Finland		France		Italy		Sweden	
Interval	$\hat{\beta}$	Interval	$\hat{\beta}$	Interval	$\hat{\beta}$	Interval	$\hat{\beta}$
5-12	1.900	9-13	3.679	5-12	3.553	10-16	1.353
13-15	0.900	14-25	0.300	13-17	0.380	17-22	0.500
16-19	0.400	26-38	0.800	18-20	0.200	23-25	0.693
20-30	0.200	39-41	0.611	21-27	0.200	26-32	0.205
31-41	0.841	42-45	0.852	28-33	0.700	33-35	0.400
42-45	0.500	46-49	0.144	34-39	0.735	36-46	0.869
46-52	0.774	50-52	0.354	40-43	1.348	47-52	0.588
-	-	-	-	44-46	0.789	-	-
-	-	-	-	47-52	0.341	-	-

In the following subsections, figures of the estimated SIR-model are presented. In addition, the policies of the ECDC government response data set are listed for each country.

4.1 Finland

Figure 1 shows the COVID19 epidemic in Finland visualized alongside with the estimated infected proportion of the population. Figure 2 is a list of policies implemented in Finland. Note that the structure of the data set containing this list may exclude or simplify some policies.

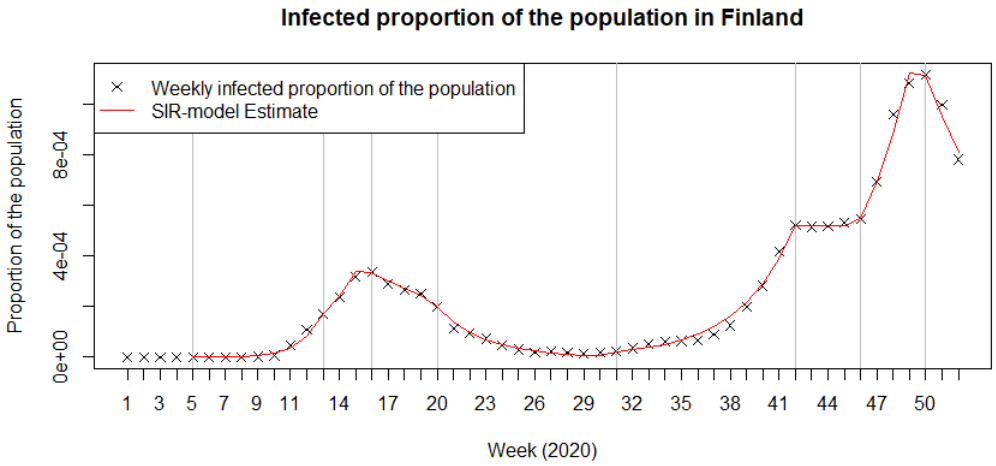

Figure 1: Weekly Finnish COVID19 infected proportion of population data alongside with the SIR-model estimate. The chosen time intervals are marked with gray vertical lines for clarity.

Response_measure	date_start	date_end
IndoorOver500	11	12
MassGatherAll	11	27
OutdoorOver500	11	12
StayHomeRiskG	12	26
ClosHigh	12	32
ClosPrim	12	20
ClosPubAny	12	22
ClosSec	12	32
EntertainmentVenues	12	22
GymsSportsCentres	12	23
IndoorOver50	12	31
MassGather50	12	27
OutdoorOver50	12	27
QuarantineForInternationalTravellers	12	24
RestaurantsCafes	12	22
Teleworking	12	31
ClosPubAnyPartial	23	NA
EntertainmentVenuesPartial	23	NA
RestaurantsCafesPartial	23	29
QuarantineForInternationalTravellersPartial	25	32
StayHomeRiskGPartial	26	NA
MassGather50Partial	27	31
MassGatherAllPartial	27	NA
OutdoorOver500	27	42
IndoorOver500	31	42
QuarantineForInternationalTravellers	32	NA
MasksVoluntaryClosedSpacesPartial	33	38
TeleworkingPartial	33	NA
MasksVoluntaryClosedSpaces	38	NA
RestaurantsCafesPartial	40	15
IndoorOver50	42	NA
MassGather50Partial	42	NA
OutdoorOver50	42	NA

Figure 2: Finnish government policies listed in the ECDC government response data set, where NA entries are in still in effect at the time of writing.

During the weeks 11-12 the first policies(Figure 2) are implemented After week 16 weekly infections decrease and remain stable between weeks 25-30. During this time a lot of policies are removed (weeks 22-27). Then, after week 31 weekly infections increase but starts to stagnate during weeks 42-45. This is again followed by an increase in growth until it peaks during week 49, followed up by a decrease again.

4.2 France

Figure 3 shows the COVID19 epidemic in France visualized alongside with the estimated infected proportion of the population. Figure 4 is a list of policies implemented in France. Note that the structure of the data set containing this list may exclude or simplify some policies.

In France, the first policies were implemented during the weeks 9-12 and sustained until the weeks 19-28. The weekly infections decrease during week 15 and remains

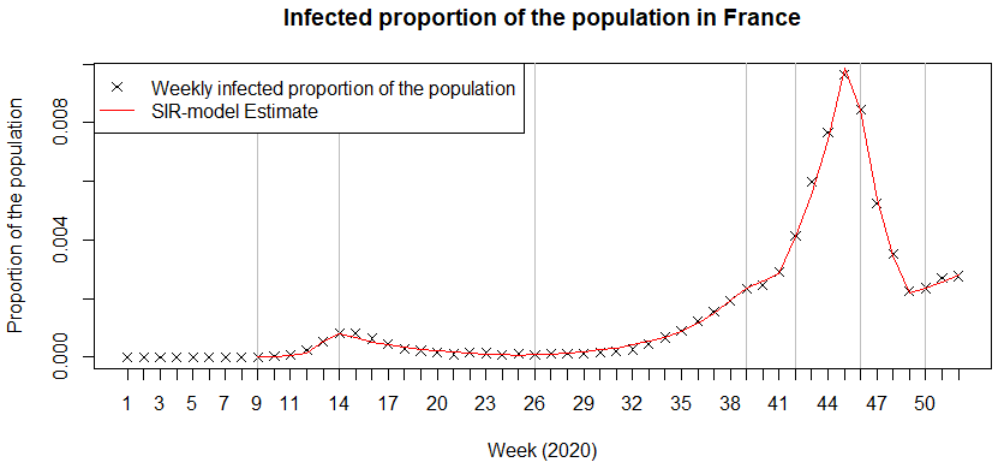

Figure 3: Weekly French COVID19 infected proportion of population data alongside with the SIR-model estimate. The chosen intervals are marked with gray vertical lines for clarity.

Response_measure	start	end
IndoorOver1000	9	42
MassGatherAll	9	38
OutdoorOver1000	11	28
ClosDaycare	12	23
ClosHighPartial	12	26
ClosPrimPartial	12	26
ClosPubAny	12	19
ClosSec	12	19
EntertainmentVenues	12	23
GymsSportsCentres	12	26
NonEssentialShops	12	19
RestaurantsCafes	12	23
PrivateGatheringRestrictions	12	25
StayHomeOrder	12	20
Teleworking	12	19
WorkplaceClosures	12	20
ClosPubAnyPartial	20	26
ClosSecPartial	20	26
MasksVoluntaryAllSpaces	20	30
MasksVoluntaryClosedSpaces	20	20
NonEssentialShopsPartial	20	23
TeleworkingPartial	20	44
StayHomeOrderPartial	20	23
MasksMandatoryClosedSpaces	21	NA
ClosDaycarePartial	23	26
EntertainmentVenuesPartial	23	26
RestaurantsCafesPartial	23	25
MassGather50Partial	38	42
MassGatherAllPartial	38	42
OutdoorOver50	38	42
PrivateGatheringRestrictionsPartial	39	44
RegionalStayHomeOrderPartial	42	44
BanOnAllEvents	43	NA
MassGather50	43	NA
MassGatherAll	43	NA
ClosHigh	44	03
ClosPubAnyPartial	44	NA
EntertainmentVenues	44	20
GymsSportsCentres	44	20
MasksMandatoryAllSpaces	44	NA
PlaceOfWorshipPartial	44	NA
PrivateGatheringRestrictions	44	NA
RestaurantsCafes	44	20
StayHomeOrder	44	51
NonEssentialShops	44	48
Teleworking	44	NA
NonEssentialShopsPartial	48	20
StayHomeOrderPartial	51	NA

Figure 4: French government policies listed in the ECDC government response data set, where NA entries are in still in effect at the time of writing.

somewhat constant between weeks 18-30. The weekly infections start to increase after week 31 then peaking during week 45. During the weeks 42-44 and 48 more strict policies are reimplemented. This seems to be followed by a rapid decrease in weekly infections after week 45.

4.3 Italy

Figure 5 shows the COVID19 epidemic in Italy visualized alongside with the estimated infected proportion of the population. Figure 6 is a list of policies implemented in Italy. Note that the structure of the data set containing this list may exclude or

simplify some policies.

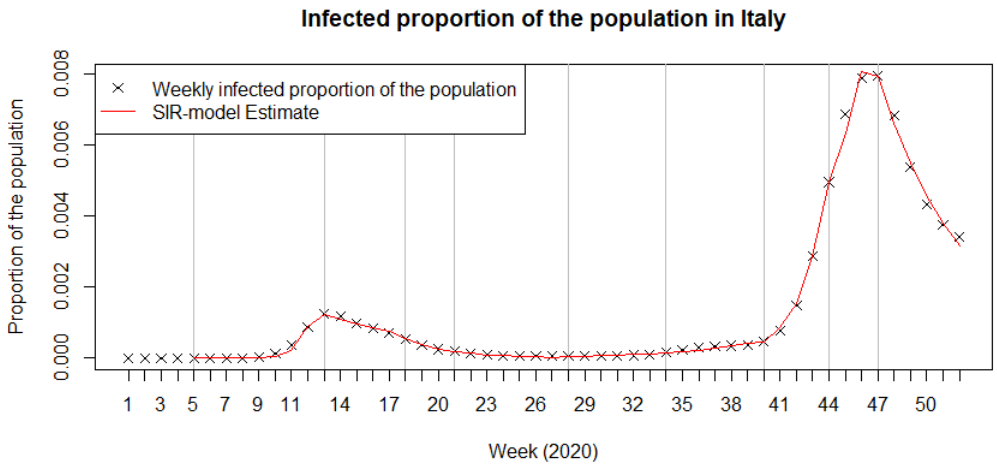

Figure 5: Weekly Italian COVID19 infected proportion of population data alongside with the SIR-model estimate. The chosen intervals are marked with gray vertical lines for clarity.

Response_measure	start	end
BanOnAllEventsPartial	10	10
ClosDaycarePartial	10	11
ClosHighPartial	10	11
ClosPrimPartial	10	11
ClosSec	10	25
MassGather50Partial	10	10
MassGatherAllPartial	10	10
QuarantineForInternationalTravellers	10	11
StayHomeRiskG	10	NA
BanOnAllEvents	11	24
MassGather50	11	24
MassGatherAll	11	24
AdaptationOfWorkplace	11	24
ClosDaycare	11	24
ClosHigh	11	24
ClosPubAnyPartial	11	52
EntertainmentVenues	11	20
GymsSportsCentres	11	21
NonEssentialShops	11	16
PlaceOfWorship	11	25
RestaurantsCafesPartial	11	21
StayHomeOrder	11	19
ClosPrim	11	38
Teleworking	11	24
NonEssentialShopsPartial	16	21
MasksMandatoryAllSpaces	17	29
EntertainmentVenuesPartial	21	25
GymsSportsCentresPartial	22	43
AdaptationOfWorkplacePartial	24	42
ClosDaycarePartial	24	38
ClosHighPartial	24	38
MassGatherAllPartial	24	45
OutdoorOver1000	24	45
QuarantineForInternationalTravellers	24	17
ClosSecPartial	25	27
ClosSec	27	38
MasksMandatoryClosedSpaces	29	33
IndoorOver500	32	45
MasksMandatoryAllSpacesPartial	33	41
MasksMandatoryAllSpaces	41	NA
Teleworking	42	NA
RestaurantsCafesPartial	42	52
RegionalStayHomeOrderPartial	43	23
EntertainmentVenues	44	12
GymsSportsCentres	44	19
StayHomeGen	44	45
ClosHighPartial	45	49
ClosSec	45	01
ClosureOfPublicTransportPartial	45	49
StayHomeOrderPartial	45	NA
BanOnAllEvents	45	16
MassGather50	45	16
MassGatherAll	45	NA
ClosHigh	49	02
ClosPubAny	52	01
NonEssentialShops	52	01
RestaurantsCafes	52	01

Figure 6: Italian government policies listed in the ECDC government response data set, where NA entries are in still in effect at the time of writing.

Policies in Italy (Figure 6) are implemented during the weeks 10-11. The weekly infections decrease initially during the weeks 13-21 and remains somewhat constant until week 34. During the weeks 24-25 many of the implemented policies are changed into partial versions or even removed. After the weeks 24-25 the weekly infections slowly start to increase. During weeks 38-39 further removal or change of policies

occurs and the increase in weekly infections accelerates until a peak is reached week 48. During weeks 43-44 the policies are reimplemented and a rapid decrease in weekly infections occurs.

4.4 Sweden

Figure 7 shows the COVID19 epidemic in Sweden visualized alongside with the estimated infected proportion of the population. Figure 8 is a list of policies implemented in Sweden. Note that the structure of the data set containing this list may exclude or simplify some policies.

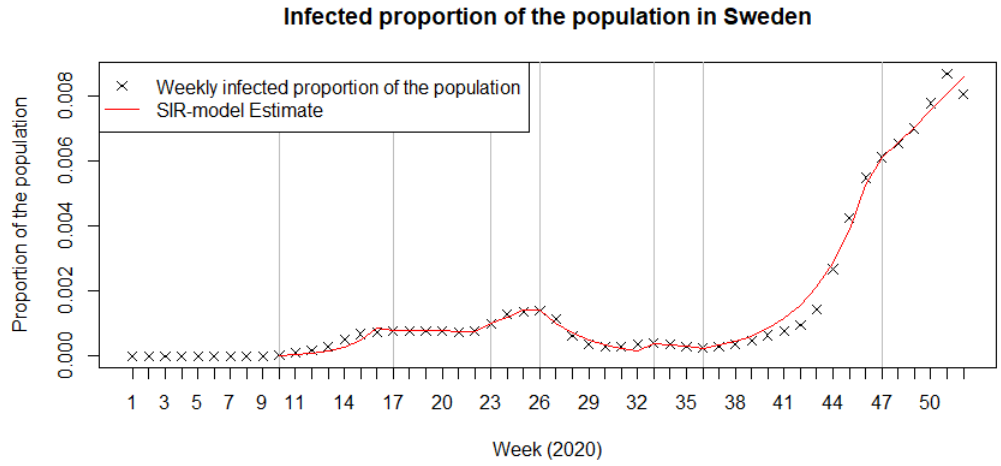

Figure 7: Weekly Swedish COVID19 infected proportion of population data alongside with the SIR-model estimate. The chosen intervals are marked with gray vertical lines for clarity.

Response_measure	date_start	date_end				
IndoorOver500	11	13		BanOnAllEventsPartial	24	09
MassGatherAll	11	24		MassGather50Partial	24	09
Teleworking	12	NA		MassGatherAllPartial	24	09
ClosHigh	12	22		ClosPubAnyPartial	27	NA
ClosSec	12	25		RestaurantsCafesPartial	27	NA
AdaptationOfWorkpl	13	28		AdaptationOfWorkplace	28	NA
BanOnAllEvents	13	24		ClosSecPartial	48	49
MassGather50	13	24		PrivateGatheringRestrictions	49	NA
StayHomeRiskG	14	43		ClosSec	50	13

Figure 8: Swedish government policies listed in the ECDC government response data set, where NA entries are in still in effect at the time of writing.

Some policies (Figure 8) are implemented in sweden during weeks 11-14, but notably less policies are listed than in the other countries.The weekly infections increase in the beginning but then stagnate during the weeks 17-21. During the weeks 27-28 more policies are implemented. A decrease in weekly infections occurs after week 26. This is then followed by stagnation in weekly infections until week 36 when weekly infections start to increase. The weekly infections slow down but are still increasing during the weeks 48-52.

5 Discussion

Based on the results presented in Section 4, a lot of similarities can be observed between the countries. The countries seem to reach similar infected proportions (Figures 1, 3, 5, 7) with the exception of Finland. An explanation for this could be that perhaps the policies were implemented fast enough to contain the epidemic. Cultural differences could be another explanation or also that government policies need not be the only reason for a reduction in the amount of new infections. For instance, public knowledge and fear of the virus might alter the behaviour of the population such that the amount of new infections reduces. In each of the chosen countries, the growth of the infected proportion of the population behaves in a somewhat cyclical manner. Starting with an exponential growth phase, a peak is reached. Then stagnation follows, until the next exponential growth phase starts. The growth phase seems to stagnate whenever policies are introduced, which is also seen in the estimated rate of transmission (Table 1). For example, when the countries implement policies during weeks 10-14 a decrease in the rate of transmission is observed. We also observe that the rate of transmission starts to increase whenever policies are removed or changed into milder versions, which seems to occur during weeks 30-40.

Based on the estimation results and the ECDC government response data sets (Figures 2, 4, 6, 8) it does seem that in general policies that reduce physical contacts, such as lockdowns, are the most effective at reducing the amount of new infections and the rate of transmission. Though, it should again be noted that because of the ECDC table format, some policies with high impact may be excluded. More complicated policies, such as targeted lockdowns are predicted to have high impact (Acemoglu et al., 2020). For example, regional lockdowns could have substantial impact on the pandemic (Roux et al., 2021), (Rossa et al., 2020).

Testing rates were not accounted for estimation, which may explain some weirdness. For instance, there are no recorded infections in the Swedish data set before week 10. But around week 10, Sweden has a larger infected proportion of the population with a smaller rate of transmission than Finland, while Finland has a lower infected proportion of the population.

The economic situations in the countries have not been taken into consideration in this thesis, which is a major limitation. If harsh restrictive policies reduce the rate of transmission, the COVID19 pandemic could be trivially stopped by having the whole population quarantined. However, such policies may be harmful to the economy (Ferguson et al., 2020), (Eichenbaum et al., 2020). The SIR-model has also been previously used to predict economic outcomes of policies. For example, Toda (2020) combined the SIR-model with an asset pricing model to predict the impact of policies on stock prices.

The SIR-model seems to fit well to the data. At least, with the assumption that the rate of transmission changes during different time intervals. This further suggests that having the rate of transmission vary not only for certain time intervals but also during every point of time improves accuracy of the fit. While the estimations in this thesis are used for comparison, the SIR-model can also be used for forecasting the

COVID19 epidemic. This has been previously explored, e.g., in Batista (2020) where a SIR-model was used for COVID19 forecasting. Forecasts using the SIR-model used in this thesis can be done by estimating the rate of transmission every time a new data entry is available. A forecast can then be calculated recursively with the SIR-model (4)-(6), using the latest data as the initial state.

6 Conclusion

In this thesis, we assesed NPI policies for COVID19 containment strategy. This was done using a SIR-model. The SIR-model was fit using sum of least squares estimation with data from four different countries. This allowed comparison between policies that reduce the spread of the COVID19 disease. Furthermore, an inspection of these policies was made using a government response data set, containing a list of policies implemented in the countries.

There were some limitations with our approach. The government response data set is really general and perhaps too simple, which may exclude policies with substantial impact, such as regional lockdowns. Most of the listed policies are also implemented at the same time. This makes it hard if not impossible to determine the impact of a single policy using the government response data set. Another problem was lack of consideration for the testing rates, which may have a substantial effect on estimation. The economic impact of these policies was also not considered. Decision makers need to assess the economic trade-offs of more harsher policies.

Overall, based on our findings and with the limitations in mind we conclude that a SIR-model can be fitted well to COVID19 data. In addition, it was also observed that for more accurate fittings the rate of transmission should vary with time. Our results also hint that NPI policies do reduce the rate of transmission. In particular the seemingly most effective policies are those that minimize physical contacts. The SIR-model used in this thesis did not account for vaccination, which is currently ongoing and could be accounted for to improve the model. For similar application of the SIR-model in the future, we suggest consideration for vaccination and economic trade-offs. Finally, for assessment of policies, we recommend the use of more thorough documentation of implemented policies.

References

Acemoglu, D., Chernozhukov, V., Werning, I. and Whinston, M. (2020), 'Optimal targeted lockdowns in a multi-group sir model', *National Bureau of Economic Research WORKING PAPER SERIES, Working Paper 27102* . DOI: 10.3386/w27102.

Batista, M. (2020), 'Estimation of the final size of the coronavirus epidemic by the sir model'.
URL: *https://www.researchgate.net/publication/339311383*

Bertozzi, A., Franco, E., Mohler, G., Short, M. and Sledge, D. (2020), 'The challenges of modeling and forecasting the spread of covid-19', *Proceedings of the National Academy of Sciences, 117 (29) 16732-16738* . DOI: 10.1073/pnas.2006520117.

Brauer, F. (2017), 'Mathematical epidemiology: Past, present, and future', *Infectious Disease Modelling, 2 (2) 113-127* . DOI: 10.1016/j.idm.2017.02.001.

Cooper, I., Mondal, A. and Antonopoulos, C. (2020), 'A sir model assumption for the spread of covid-19 in different communities', *Chaos, Solitons Fractals, 139 110057, ISSN 0960-0779* . DOI: 10.1016/j.chaos.2020.110057.

ECDC (2021*a*), 'Data on 14-day notification rate of new covid-19 cases and deaths', *European Centre for Disease Prevention and Control* .
URL: *https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19*

ECDC (2021*b*), 'Data on country response measures to covid-19', *European Centre for Disease Prevention and Control* .
URL: *https://www.ecdc.europa.eu/en/publications-data/download-data-response-measures-covid-19*

Eichenbaum, M., Rebelo, S. and Trabandt, M. (2020), 'The macroeconomics of epidemics', *National Bureau of Economic Research WORKING PAPER SERIES, Working Paper 26882* . DOI: 10.3386/w26882.

Ferguson, N., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., Dighe, A., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L., van Elsland, S., Thompson, H., Verity, R., Volz, E., Wang, H., Wang, Y., Walker, P., Walters, C., Winskill, P., Whittaker, C., Donnelly, C., Riley, S. and Ghani, A. (2020), 'Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand', *Imperial College COVID-19 Response Team* . DOI: 10.25561/77482.

Kermack, W. and McKendrick, A. (1927), 'A contribution to the mathematical theory of epidemics', *Proceedings of the Royal Society* . DOI: 10.1098/rspa.1927.0118.

Mickens, R. (1988), 'Difference equation models of differential equations', *Mathematical and Computer Modelling, 11 528-530, ISSN 0895-7177* . DOI: 10.1016/0895-7177(88)90549-3.

Rossa, D., Salzano, D., Di Meglio, A., Coraggio, M., Calabrese, C., Guarino, A., Cardona-Rivera, R., De Lellis, P., Liuzza, D., Lo Iudice, F., Russo, G. and di Bernardo, M. (2020), 'A network model of italy shows that intermittent regional strategies can alleviate the covid-19 epidemic', *Nature Communications 11 (5106)* . DOI:10.1038/s41467-020-18827-5.

Roux, J., Massonnaud, C., Colizza, V., Cauchemez, S. and Crépey, P. (2021), 'Impact of national and regional lockdowns on covid-19 epidemic waves: Application to the 2020 spring wave in france', *medRxiv* . DOI:10.1101/2021.04.21.21255876.
URL: *https://www.medrxiv.org/content/early/2021/04/25/2021.04.21.21255876*

Toda, A. (2020), 'Susceptible-infected-recovered (sir) dynamics of covid-19 and economic impact'.
URL: *https://arxiv.org/abs/2003.11221*

You, C., Deng, Y., Hu, W., Sun, J., Lin, Q., Zhou, F., Heng Pang, C., Zhang, Y., Chen, Z. and Zhou, X.-H. (2020), 'Estimation of the time-varying reproduction number of covid-19 outbreak in china', *International Journal of Hygiene and Environmental Health, 228 113555, July 2020, ISSN 1438-4639* . DOI:10.1016/j.ijheh.2020.113555.

Appendix A R code

```
getData<-function(country,x0,xfin){
  serialtime = 2 # (infectious time) 2 weeks forall

  ecdc.data <- read.csv(file="ecdc1.csv",header=TRUE)

  ecdc.data<- subset(ecdc.data, ecdc.data$ï..country==country)
  ecdc.data<-subset(ecdc.data, ecdc.data$indicator=="cases")

  N<-median(ecdc.data$population)
  dI<-ecdc.data$weekly_count[1:52]

  myint<-x0:xfin

  C<-1:length(dI) #calculate cumulative cases
  for(i in 1:length(dI)){
    C[i]<-sum(dI[1:i])
  }

  dR<-1:length(dI) #calculate daily recoveries
  for(i in 1:length(dI)){
    if(i-serialtime<=0){
      dR[i]=0
    }else{
      dR[i]<-dI[i-serialtime]
    }
  }

  R<-dR #calculate recovered population
  for(i in 1:length(dR)){
    R[i]<-sum(dR[1:i])
  }

  S<-N-C
  I<-C-R

  rm(i) # i was used in for loops, removed in case of errors
  #normalize
  s<-S[myint]/N
  i<-I[myint]/N
  r<-R[myint]/N
  return(cbind(s,i,r,N))
}
```

```
sir <- function(beta,gamma,s,i,r){
  #calculate the SIR variables, starting from initial states of
  #s,i,r
  x<-s
  y<-i
  z<-r

  for(t in 2:length(i)){
    x[t]=x[t-1]-beta*x[t-1]*y[t-1]
    y[t]=y[t-1]+beta*x[t-1]*y[t-1]-gamma*y[t-1]
    z[t]=z[t-1]+gamma*y[t-1]
  }
  return(list(x,y,z))
}

loss <- function(S,s){
  l = sum( (S-s)^2 )
  return(l)
}

estimateSIR<-function(s,i,r){

  k=0             #iteration count
  stop=F          #stops the loop
  g=1/2           #gamma, 1/2weeks

  bub = 7         #upper and lower limits of beta
  blb = 0


  b=1             #initial beta guess

  l=100           #high initial loss function value

  h=1/10000       #a really small number
  stepsize = 1/10 #default stepsize
  maxiter=1000    #max iterations
  myls = 1:maxiter #store l
  mybs = 1:maxiter #store betas
  improvelimit<-h^6 #small number

  while(k<maxiter & !stop){
    sir.temp <- sir(b,g,s,i,r)
    i.e<-unlist(sir.temp[2])
```

```
l.prev<-l
l<-loss(i,i.e)

if(log(l)< -20){ #stop if l small
  stop=T
  message("Done, b=",b)
  break
}

if(abs(l-l.prev)< improvelimit){ #stop if nothing improves
  stop=T
  message("Cant improve, b=",b)
  break
}

step=stepsize
stepcounter=0
while(b+step>bub | b-step<blb){
  #modify stepsize to stay
  #within bounds
  step=step/10
  stepcounter=stepcounter+1
  if(stepcounter>100){
    stop=TRUE
    message("Out of bounds")
    break
  }
}
stepcounter=0

#using + or - stepsize calculate the i variable
sir.p=sir(b+step,g,s,i,r)
sir.m=sir(b-step,g,s,i,r)

#then loss function value
l.plus<- loss(i, (unlist(sir.p[2])))
l.minus<- loss(i, (unlist(sir.m[2])))
while(l.plus>l & l.minus>l){
  #make sure the next point is within bounds

  sir.p=sir(b+step,g,s,i,r)
  sir.m=sir(b-step,g,s,i,r)
```

```
      stepcounter=stepcounter+1
      step=step/10
      l.plus<- loss(i, (unlist(sir.p[2])))
      l.minus<- loss(i, (unlist(sir.m[2])))

      if(stepcounter>30){
        message("error")
        stop=T
        break
      }
    }

    k=k+1
    #store betas and l for debugging
    mybs[k]<-b
    myls[k]<-l

    if(l.plus>=l.minus){
      b=b-step
    }else{
      b=b+step
    }

  }
  mybs<-mybs[1:k]
  myls<-myls[1:k]


  #return variables as a matrix
  sir.temp<-sir(b,g,s,i,r)
  s.e<-unlist(sir.temp[1])
  i.e<-unlist(sir.temp[2])
  r.e<-unlist(sir.temp[3])

  return(cbind(s.e,i.e,r.e,b))
}


estimateCountry <- function(country,intervals){
  #run estimation for a specific country
  #produces plots and the i variable of that country
  #as well as betas list


  x0=1
```

```
  xfin=52

  SIR<-getData(country,x0,xfin)
  i<-SIR[,2]
  N<-SIR[1,4]


  intslen<-length(intervals)
  bs<- NA
  is<-NA


  for(j in 2:(unlist(intervals[1])[1]-1)){
    is<-append(is,NA)
  }

  for(j in (intervals)){
    x0<-unlist(j)[1]
    xfin<-unlist(j)[length(j)]
    SIR<-getData(country,x0,xfin)
    estimate<-estimateSIR(SIR[,1],SIR[,2],SIR[,3])


    bs<-append(bs,estimate[1,4])
    is<-append(is,estimate[,2])

  }


  plot(i, xaxt="n", main=paste("Infected proportion of the population in"
                                  ,country),xlab="Week (2020)",pch=4)
  for(j in 1:length(intervals)){
    abline(v=unlist(intervals[j])[1], col="gray")
  }

  lines(is,col="red", lwd=1)
  legend("topleft",legend=c("Original", "Estimate")
         ,col=c("black","red"), pch=c(4,NA),lwd=c(NA,1))
  axis(side=1,at=seq(from=1,to=52,by=1))

  return(is)
}

#Store interval list for estimations
```

```
#ITA
iintervals<-list(5:12,
                 13:17,
                 18:20,
                 21:27,
                 28:33,
                 34:39,
                 40:43,
                 44:46,
                 47:52
)



#FIN
fintervals<-list(
  5:12,
  13:15,
  16:19,
  20:30,
  31:41,
  42:45,
  46:49,
  50:52
)

#FRA
frintervals<-list(
  9:13,
  14:25,
  26:38,
  39:41,
  42:45,
  46:49,
  50:52
)

#SWE
sintervals<-list(
  10:16,
  17:22,
  23:25,
  26:32,
  33:35,
  36:46,
```

```
  47:52
)
```

```
estimateCountry("Finland",fintervals)
estimateCountry("France",frintervals)
estimateCountry("Italy", iintervals)
estimateCountry("Sweden", sintervals)
```